# Implementación de Blockchain para aseguramiento de evidencia digital en entornos Forensic Readiness

Javier Díaz[1], Mónica D. Tugnarelli[2], Lucas Barboza[2], Mauro F. Fornaroli[2], Facundo N. Miño[2]
[1] Facultad de Informática – Universidad Nacional de La Plata
[2] Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos
jdiaz@unlp.edu.ar, {monica.tugnarelli, lucas.barboza, mauro.fornaroli}@uner.edu.ar

**Abstract.** En este artículo se presentan los avances PID-UNER 7059 en la implementación de blockchain aplicada a la preservación de evidencia digital en entornos Forensic Readiness el cual plantea requisitos estrictos de recolección de los datos antes de que se produzca un incidente de seguridad. Se analizan las estructuras y tipos de blockchain disponibles, clasificándolas según diversos parámetros. En base a ese análisis se seleccionaron dos, Ethereum bajo especificaciones de uso de la Blockchain Federal Argentina como blockchain pública y la segunda, del tipo blockchain privada con el framework Hyperledger Fabric. Ambas se utilizarán como entornos de prueba de concepto, dentro de un esquema genérico de funcionamiento y con determinadas restricciones considerando el tipo de datos que se debe resguardar y las características que demanda la cadena de custodia de este tipo de evidencia.

**Keywords:** blockchain, forensic readiness, evidencia digital, cadena de custodia

## 1 Introducción

En este trabajo se presentan los avances del PID-UNER 7059, iniciado en febrero de este año, denominado "*Tecnología Blockchain para aseguramiento de evidencia digital en entornos Forensic Readiness*" que tiene como objetivo principal analizar el impacto de la utilización de la tecnología blockchain aplicada a la preservación, integridad y trazabilidad de evidencia digital, obtenida de activos esenciales de una organización, en un entorno preventivo como lo es Forensic Readiness, también llamado Preparación Forense.

Específicamente, se pretende disponer de un prototipo de blockchain sin criptomoneda asociada, sobre el cual se puedan realizar pruebas que permitan analizar las prestaciones frente a las características demandadas por Forensic Readiness para el proceso de asegurar la evidencia y mantenimiento de la cadena de custodia. Asimismo se busca lograr una adecuada integración de la blockchain con esquemas de recolección de datos con las características que impone la Preparación Forense.

Forensic Readiness propone que la evidencia digital se recolecte y asegure de manera anticipada, es decir, antes de la ocurrencia de un incidente de seguridad. Este término fue enunciado por John Tan [1] quien lo describió principalmente a través de dos

objetivos: maximizar la capacidad del entorno para reunir evidencia digital confiable y minimizar el costo forense durante la respuesta a un incidente.

En este enfoque, que fue analizado en artículos anteriores [2] [3] [4], los datos que se recolectan pueden ser utilizados como insumo para el análisis de incidentes de seguridad y también como prueba legal, lo que involucra el aseguramiento de la prueba a medida que se realiza la recolección activa de los datos, tarea que fue realizada anteriormente utilizando funciones hash para resguardar la integridad de la evidencia digital. Considerando estos requisitos, una instancia fundamental para garantizar su admisibilidad como elemento de prueba es la preservación de la Cadena de Custodia. Esta Cadena de Custodia debe estar claramente documentada y con un registro detallado desde su recolección hasta su almacenamiento, por lo que la aplicación de la tecnología Blockchain resulta de especial interés para cumplimentar esta función.

Por su parte, Blockchain puede describirse como una base de datos distribuida y organizada bajo una estructura de conjunto de bloques que se van encadenando entre sí mediante dos códigos hash que actúan como enlaces: uno para el bloque de datos creado anteriormente y otro para que se comparta y grabe en el bloque que se cree a continuación, de forma tal de obtener una lista enlazada o cadena de bloques. Cada bloque que se encadena se vuelve inmutable y en eso radica la fortaleza de seguridad y verificabilidad de esta tecnología en la que, cuanto mayor es el grado de replicación de los bloques, más sencillo resulta detectar adulteraciones. [5].

Las criptomonedas fueron una de las primeras aplicaciones visibles que tienen como base esta tecnología, pero más allá de ellas, blockchain también permite imaginar múltiples prestaciones relacionadas a transacciones, a la seguridad, a la trazabilidad y a la transparencia.

Como primera etapa del proyecto se relevaron casos de uso a nivel regional y nacional [6], entre las que se destaca la Blockchain Federal Argentina [7] que brinda una plataforma pública para integrar servicios y aplicaciones sobre blockchain siguiendo el modelo de Múltiples Partes Interesadas. BFA mantiene un modelo de gobernanza que asegura la representación de todos los sectores en la toma de decisiones pero, al ser una plataforma pública, su uso no está restringido solo a las organizaciones que integran el consorcio.

Entre los participantes de esta plataforma se encuentra la Facultad de Ciencias de la Administración como miembro parte y encargado del nodo sellador UNER, y dos integrantes de este proyecto, en carácter de responsables designados ante la BFA en representación de la mencionada institución.

## 2 Análisis de estructuras y tipos de blockchain

La segunda etapa del PID se focaliza en el análisis de distintos tipos de blockchain para contar con una base de conocimiento previa a la implementación de los prototipos para la realización de pruebas. A continuación se presenta un breve resumen de esta revisión.

De manera general, las redes blockchain pueden clasificarse como [8], [9], [10], [11]:

- **Públicas:** donde cualquiera puede acceder a los datos de la cadena de bloques. Normalmente son transparentes, los usuarios son anónimos, no existe un administrador de la red y las transacciones se validan mediante un protocolo de consenso.
- **Privadas:** son aquellas donde existe una entidad central que se encarga de controlar la cadena de bloques, definir la lista de participantes autorizados, otorgar permisos, proponer transacciones y validar los bloques. Puede decirse que no existe descentralización ni consenso ya que es una única entidad quien administra la red. Los usuarios finales dependen de una interfaz provista por el administrador para leer o enviar transacciones.
- **De Consorcio o Federadas**, son las que tienen un conjunto de participantes predefinido, tales como empresas u organizaciones, quienes se encargan de la administración conjunta de la red y de asegurar el mantenimiento sincronizado de las copias del registro compartido. El acceso a los datos puede ser público o privado. Son redes parcialmente descentralizadas, adecuadas para aplicaciones en donde se generan grandes volúmenes de transacciones entre entidades con requerimientos de confianza mutua.

Otra clasificación posible surge de considerar quiénes poseen permisos para crear bloques, a saber:

- **Sin permisos (*permissionless*)**: cualquier entidad puede procesar transacciones, participar de los protocolos de consenso y crear bloques.

- **Con permisos (*permissioned)*:** solo una lista o conjunto de entidades predefinidas, y con identidades conocidas, pueden participar del procesamiento de transacciones, lo que agrega una capa más para el control de acceso e identificación.

En definitiva, las redes públicas pueden o no tener permisos, mientras que las privadas y las de consorcio o federadas suelen ser de tipo permisionadas. Las aplicaciones de blockchain con criptomonedas, como Bitcoin o Ethereum, son ejemplos de redes públicas sin permisos, mientras que el proyecto Hyperledger permite implementar redes de blockchain privadas con permisos.

Por último, desde el punto de vista de la realización de las transacciones, se puede hacer una diferenciación entre off-chain y on-chain, donde:

- **Transacciones on-chain (dentro de la cadena)**: son las que ocurren dentro del blockchain, adquieren validez sólo cuando la cadena de bloques se

modifica para registrar la transacción en el ledger distribuido y son visibles para todos los nodos participantes. El tiempo de esta operación depende del volumen de transacciones y la cantidad de nodos selladores que actúan en la red. Por lo general los mineros cobran una comisión por sus servicios de validación y autenticación de transacciones.

- **Transacciones off-chain**: en este tipo de transacciones los datos en cuestión permanecen fuera de la cadena de bloques, quedando dentro de la blockchain el ID y el hash que la identifica. En contraste con las transacciones on-chain, los datos no son de acceso público y las transacciones se ejecutan de manera inmediata. Implican un acuerdo entre las partes para realizar la transacción y, si fuera necesario, la participación de un tercero para validar la operación.

### 2.1. Ethereum versus Hyperledger

De acuerdo a los objetivos de esta investigación se han analizado dos soluciones representativas de blockchain, una pública y otra privada, considerando para esto prestaciones tales como: privacidad, seguridad, velocidad de validación de transacciones, casos de uso, estándar abierto, entre otros ítems. En este sentido:

- **Ethereum** es una plataforma pública, distribuida y descentralizada que funciona en una red virtual llamada Ethereum Virtual Machine (EVM) donde cada nodo participante paga una comisión por cada transacción realizada utilizando para ello la criptomoneda Ether. Cuenta con una billetera, herramientas de línea de comandos, una aplicación GUI y contratos inteligentes (Smart Contract) de uso público y descentralizado para acuerdos entre partes. Como es una red pública no requiere de permisos y es completamente transparente, lo que significa que todas las transacciones registradas en la red blockchain son visibles y accesibles para todos los nodos participantes los cuales implementan algoritmos de consensos llamados Prueba de Trabajo (PoW) para aceptar cada transacción. Con Ethereum, cualquiera puede crear un nodo y cada nodo en la red poseerá una copia de la Blockchain. La red posee una infraestructura de nodos a nivel global. Como el desarrollo está basado en código abierto, toda la comunidad puede participar en las pruebas de concepto existentes para mejorar la plataforma, o tomar todo ese trabajo y adaptarlo a otros contextos y necesidades [12].

En este proyecto se utilizará la infraestructura disponible de BFA que toma el software de Ethereum, utilizando Prueba de Autoridad (PoA), sin criptomoneda asociada.

- Por su parte, **Hyperledger** es un proyecto de código abierto de la Linux Foundation, con la colaboración de empresas y organizaciones, con la finalidad de desarrollar soluciones con frameworks DLTs (Distributed Ledger Technology) conocidos también como blockchains privadas. Cuenta además con motores de smart contracts, librerías clientes e

interfaces gráficas. Tiene un diseño modular, extensible e interoperable que permite identificación de participantes, privacidad y confidencialidad de las operaciones, alto rendimiento para carga y validación de las transacciones y administración centralizada. El framework Hyperledger Fabric demuestra un alto rendimiento en términos de procesamiento de transacciones y baja latencia de confirmación de las mismas, además de permitir la privacidad y confidencialidad de las transacciones y el uso de contratos inteligentes. [13] [14] [15]. En relación al objetivo de investigación planteado se destacan las siguientes características:

- Es una red permisionada privada, con acceso restringido y la identidad de los participantes es conocida.
- El consenso se realiza a nivel transaccional, lo que significa que no es necesario validar todo el bloque lo que aumenta la velocidad de las transacciones.
- No requiere de una criptomoneda para su funcionamiento y prescinde del protocolo de consenso PoW (*Proof of Work*) que es costoso de implementar debido al proceso de minería que requiere y, por ende, consume menos recursos.

### 2.2 Esquema de funcionamiento

En la siguiente figura se presenta un esquema genérico de funcionamiento el cual se utilizará como marco para las pruebas de laboratorio.
En la imagen se describe el proceso de almacenamiento de los hashes de la evidencia recolectada de los activos esenciales determinados, junto a los metadatos de la misma, en el ledger (libro de registro distribuido) para permitir la correlación. De esta manera se resguarda la información de la evidencia en una base de datos con mayor rapidez y seguridad, en comparación a las transacciones on-chain.

El procedimiento es el siguiente:

- Se determina el activo digital sobre el cual se realizará la recolección de datos considerados evidencia.
- Se obtiene el hash o identidad digital de esos datos que junto al sello de tiempo (timestamp) proporcionarán la integridad.
- La evidencia se almacena en una base de datos off-chain o directorio privado
- El hash y los metadatos de la evidencia se registran en la blockchain.

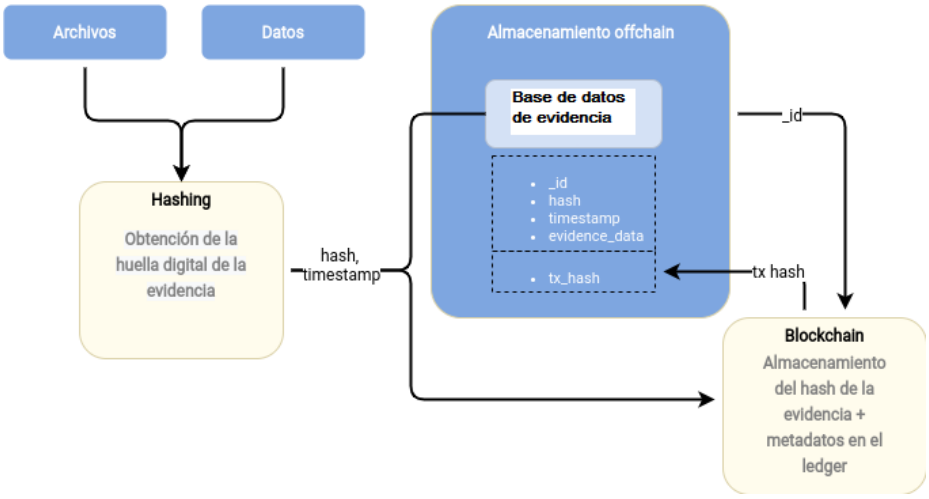


Fig.1 Almacenamiento de evidencia digital en Blockchain con preservación off-chain

Algunos puntos a considerar al momento de realizar las pruebas de concepto:

a. En vista de los requisitos relacionados con Forensic Readiness, es necesario contar con una blockchain donde los datos, posiblemente confidenciales, puedan almacenarse fuera de la cadena de bloques.
b. El sello de tiempo es un mecanismo que permite generar una "prueba de existencia" de un archivo digital. En este sentido, BFA ofrece su propio servicio de servicio de TSA (*Time Stamping Authority*).
c. Es necesario que la base de datos que almacenará la evidencia recolectada tenga como mínimo un backup. Tanto la base como su respaldo deben ser tratados como activos esenciales de la organización. En este punto es indispensable seleccionar una infraestructura adecuada para minimizar los riesgos de borrado, pérdida o destrucción.
d. Como en Hyperledger Fabric no se necesita el consenso de todos los nodos para validar las transacciones, se utiliza un algoritmo BTF (Tolerancia a las fallas bizantinas) [16] con la posibilidad de usar más de un mecanismo de consenso para resolver los problemas de confianza entre los nodos.
e. Hyperledger Fabric ofrece mayor customización para casos de uso que requieran estricta privacidad.
f. Al ofrecer transacciones sin costo**,** BFA cuenta con una "Destilería de Gas" que regula el envío de ether a nodos transaccionales. Esto puede tener impacto y ser un limitante según el volumen de datos recolectados.

g. Dadas las implicancias de un proceso legal que requiera el uso de la evidencia almacenada, será necesario contar con una tercera parte que permita validar y asegurar la integridad de las transacciones de forma independiente. La estructura de múltiples partes interesadas de BFA asegura esa independencia. En Hyperledger Fabric puede darse a través de la incorporación de uno o más nodos de confianza a la red blockchain permisionada, los que no sólo tendrán participación en la validación y confirmación de transacciones, sino que además guardarán su propia copia de los registros.
h. La admisibilidad de una evidencia digital ante un proceso legal requiere la documentación cronológica y detallada de la Cadena de Custodia, por ende será determinante el mantenimiento de los requisitos de integridad, trazabilidad, autenticación y verificación de la misma que se alcance en cada una de las soluciones a analizar.

Para configuración del entorno de trabajo con la blockchain pública se usará la infraestructura disponible con la instalación de un nodo transaccional en la red de prueba (test2), sobre Ubuntu Server, de acuerdo a los requerimientos técnicos especificados por BFA en su Wiki.
Por otra parte también se ha configurado un entorno con Hyperledger Fabric mediante contenedores de Docker [17] para desplegar servicios en una máquina virtual de 4GB RAM, 150GB de disco HDD, 4 núcleos y Ubuntu Server como sistema operativo, sobre una infraestructura propia que utiliza XenServer (XCP-ng) [18] como plataforma de virtualización. En esta estructura mínima se han definido clientes, nodos pares y un nodo que actúa de Autoridad de Certificación para los usuarios.

## 3 Conclusiones y trabajo a futuro

En este artículo se han presentado los resultados de las primeras etapas del PID 7059, avanzando en el análisis de tipos y prestaciones de redes blockchain públicas y privadas para seleccionar la estructura que mejor se adapte a los requerimientos planteados por el entorno Forensic Readiness.

Luego de la revisión se han seleccionado por un lado el framework Hyperledger Fabric como blockchain privada y por otro una implementación de Ethereum de acceso público tal como es la Blockchain Federal Argentina con los cuales se ha instalado en un entorno virtual que será el ambiente de testing a los fines de comparar cómo responde cada una de las blockchains elegidas a la dinámica relacionada con la recolección y resguardo de evidencia digital y la implementación de la Cadena de Custodia exigida.

Asimismo, se ha presentado el esquema básico de trabajo y funcionamiento de la blockchain con la particularidad de almacenamiento con preservación off-chain de los datos.

Las actividades a futuro comprenderán para los dos casos de práctica: la especificación del procedimiento de recolección, almacenamiento y acceso a la base de evidencias, la descripción del procedimiento de autenticación y trazabilidad de la evidencia digital incluyendo la documentación de la cadena de custodia, el correspondiente análisis de performance y escalabilidad, así como también la revisión de aspectos de seguridad relacionados con ambas implementaciones.

### Referencias


[1] TAN, John. (2001). Forensic Readiness. http://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf

[2] Tugnarelli, M.; Fornaroli, M.; Santana, S.; Jacobo, E.; Díaz, F.J. Análisis de metodologías de recolección de datos digitales. Workshop de Investigadores en Ciencias de la Computación (WICC 2017). ISBN: 978-987-42-5143-5. http://sedici.unlp.edu.ar/handle/10915/61343

[3] Tugnarelli, M., Fornaroli,M. Santana,S. Jacobo, E. Díaz, J. Analysis of Methodologies of Digital Data Collection in Web Serves. Communications in Computer and Information Science (Springer), Vol. 790, Pag.265. (2018) https://link.springer.com/content/pdf/bfm%3A978-3-319-75214-3%2F1.pdf

[4] Tugnarelli, Mónica; Díaz Francisco Javier. Forensic Readiness: Guía de Buenas Prácticas. Libro de Actas. XXV Congreso Argentino de Ciencias de la Computación CACIC 2019. VI Workshop de Seguridad Informática, pp. 1261-1268. ISBN 978-987-688-377-1.

[5] Michael Crosby, et. al. BlockChain Technology: Beyond Bitcoin. Applied Innovation Review (AIR). Issue No. 2 June 2016. Berkeley. http://scet.berkeley.edu/wp-content/uploads/AIR-2016-Final-version-Int.pdf

[6] Javier Díaz, Mónica D. Tugnarelli, Mauro F. Fornaroli, Lucas Barboza. Blockchain para Aseguramiento de Evidencia Digital en entornos Forensic Readiness.Workshop de Investigadores en Ciencias de la Computación (WICC 2020). ISBN: en trámite. Exposición virtual de posters: https://drive.google.com/file/d/1vGiNX9ogumSBjnAnH-4d6N5hn5omJPM4/view

[7] Blockchain Federal Argentina https://bfa.ar

[8] Iuon-Chang Lin, Tzu-Chun Liao. A Survey of Blockchain Security Issues and Challenges . International Journal of Network Security, Vol.19, No.5, PP.653-659, Sept. 2017

[9] Elli Androulaki, Artem Barger, Vita Bortnikov, et al, 2018. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains . In EuroSys '18: Thirteenth EuroSys Conference 2018, April 23–26, 2018, Porto, Portugal. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3190508.3190538

[10] BitFury Group. Public versus Private Blockchains Part 1: Permissioned Blockchains. White Paper. Oct 20, 2015 (Version 1.0). https://bitfury.com/content/downloads/public-vs-private-pt1-1.pdf

[11] BitFury Group. Digital Assets on Public Blockchains. White Paper. Mar 15, 2016 (Version 1.0). https://bitfury.com/content/downloads/bitfury-digital_assets_on_public_blockchains-1.pdf

[12] Ethereum. https://ethereum.org

[13] Hyperledger. https://www.hyperledger.org/

[14] An Introduction to Hyperledger.V1.1. Published August 2018. https://www.hyperledger.org/wp-content/uploads/2018/08/HL_Whitepaper_IntroductiontoHyperledger.pdf

[15] Hyperledger Fabric: the flexible blockchain framework that's changing the business world https://www.ibm.com/blockchain/hyperledger

[16] Lei, Kai & Zhang, Qichao & Xu, Limei & Qi, Zhuyun. (2018). Reputation-Based Byzantine Fault-Tolerance for Consortium Blockchain. 10.1109/PADSW.2018.8644933.

[17] Docker Docs. https://docs.docker.com/

[18] XCP-ng. https://xcp-ng.org/